

SECRET

17 January 1969

MEMORANDUM FOR: COINS Project Manager

SUBJECT : Standardization in COINS Files

REFERENCES : 1) My Memo to You, 8 Jan 1969,
Same Subject
2) COINS/192, Draft COINS Operating
Procedure on Data Standards in
the COINS Experiment, 2 Jan 1969

1. In my memorandum to you dated 8 January 1969 (Reference 1), I enumerated several problems which would inhibit attempts to standardize data entries in COINS files. Receipt of your most recent memorandum (Reference 2) proposing a COINS Operating Procedure for Data Standardization prompts me to further elaborate on this subject.

2. I believe an attempt to achieve data standards in the COINS Experiment is commendable. However, I question whether on a basis of priorities it is the most critical problem we should be attacking at the moment. Right now I am most concerned with seeing the communications/software aspects of the COINS network operating in an efficient manner. I would not recommend diluting this task by diverting our programming support to standardization efforts. My next major concern is with obtaining additional data on the query languages and files maintained by other COINS participants so that I will be in a position to train CIA personnel to query those systems prior to the start of the operational test. I fear that effort devoted to standardization might delay my receiving this information.

3. I also question whether it is feasible to expect to achieve any degree of standardization in the near future, and if something cannot be achieved in the short term, then it should be deferred until

GROUP 1
Excluded from automatic
downgrading and
declassification

~~SECRET~~

LWIX

after the experiment. Anyone with experience in information storage and retrieval experiments can attest that you have to freeze the experimental design prior to start of the operational test because further change will have an adverse, unmeasurable effect on results. I think COINS users would become understandably confused and discouraged during the experiment if they suddenly discovered that certain data was no longer retrievable in a manner to which they had become accustomed or worse if they were unaware that data had been changed and their searches were negative because they did not specify the correct data item in the search prescription. The potential COINS users in CIA are not a well defined group that I could easily transmit changes to. For this reason, I am opposed to effecting changes in data standards during the operational test not only in CIA files, but also in the files of other COINS participants.

4. I am not optimistic concerning the capability of COINS participants to adopt common data standards in the near future. The COINS Managers could probably agree, after several discussion sessions, to a limited set of standards on less controversial fields such as how to enter dates. However, as you are aware, COINS Managers do not control the COINS files (at least this manager does not); therefore, negotiations are required with the file builders and others within each agency. This would be followed by renegotiation among agencies and probably coordination with the IHC. If after this procedure, common standards can be agreed upon, then each agency must find and allocate the manpower necessary to effect changes. This is a time consuming effort and I can cite numerous instances where agreed upon standards have never been applied to existing files because of the large effort that would be required to effect the changes. I personally have participated in two interagency committees that were established to develop respectively a common geo-political area code and a standard subject code. Each of these efforts took well over a year and the resultant agreed upon codes were never officially adopted by some agencies, even for new EDP files.

} You're supposed to manage them

~~SECRET~~

SECRET

L-111

5. The above discussion is not presented to imply that I am opposed to standardization. Standardization would obviously facilitate the use of COINS files. However, the tone of your proposed operating procedure implies that something is achievable now, whereas it is my opinion that if we start working on the problem now we may be able to obtain agreed upon standards for a very limited set of data items by 1970. Further, I feel that the IHC should be directly involved in any standardization efforts because if we adopt standards for COINS files, we are in effect adopting standards for a significant number of inter-agency EDP files, and these standards will continue in effect regardless of the outcome of COINS. This should be of concern to the IHC.

} Right!?

Right

6. There is one area where I think we could rather easily achieve standardization with significant benefit to the user. This area concerns field names or data elements as defined in your memorandum. This seems to me the most logical place to begin because:

- a. Users have become accustomed to dealing with different code schemes and data entry procedures among different files. However, they will not be able to comprehend why the same field in two different files is referred to with different acronyms, e.g., the surname field is referred to as NAME, SRNAM, and NAML in three different COINS files.
- b. These data element acronyms are merely labels for which no standardization precedence exists among agencies. We would not be dealing with a sensitive area involving previously established prerogatives.
- c. Changing these labels is a simple task compared to changing data items.

the same name would be used for different fields.

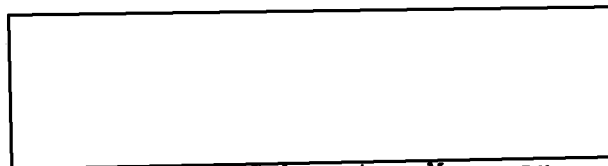
7. I recommend that we undertake this task immediately. Perhaps a member of your staff could compile a list of data element acronyms in use in COINS files with a recommended standard for those

SECRET

SECRET

fields in which multiple acronyms exist. If we cannot arrive at agreement in this area, then I would not be very hopeful concerning our ability to agree on data standards in other areas.

8. I have no objection to the basic principles enunciated in the referenced proposed Operating Procedure. I question whether it is necessary to formalize the procedure in writing, i.e., I am opposed to documentation for the sake of documentation. However, if you feel that it is desirable to publish it, I have several recommended changes which I can transmit separately.



CIA COINS Subsystem Manager

STAT

SECRET